Europäisches Patentamt

(19) European Patent Office

Office européen des brevets

(11) **EP 0 841 548 A2**

(12) **EUROPEAN PATENT APPLICATION**

(43) Date of publication:
13.05.1998 Bulletin 1998/20

(51) Int Cl.6: **G01J 11/00, G01N 21/35**

(21) Application number: 97308573.1

(22) Date of filing: 28.10.1997

(84) Designated Contracting States:
AT BE CH DE DK ES FI FR GB GR IE IT LI LU MC NL PT SE
Designated Extension States:
AL LT LV RO SI

(30) Priority: 06.11.1996 US 744484

(71) Applicant: LUCENT TECHNOLOGIES INC.
Murray Hill, New Jersey 07974-0636 (US)

(72) Inventors:
• Jacobsen, Rune Hylsberg
  8000 Aarhus (DK)
• Nuss, Martin C.
  Fair Haven, New Jersey 07704 (US)
• Mittleman, Daniel Matthew
  Houston, Texas 77025 (US)

(74) Representative:
Watts, Christopher Malcolm Kelway, Dr. et al
Lucent Technologies (UK) Ltd,
5 Mornington Road
Woodford Green Essex, IG8 0TU (GB)

(54) **Systems and methods for processing and analyzing terahertz waveforms**

(57) A time-domain signal processing system for displaying, classifying, and recognizing temporal and spectral features in terahertz waveforms returned form ; materials. Specifically, novel techniques are described for classifying and analyzing the free induction decay exhibited by gases excited by far-infrared (terahertz) pulses. Illustratively, a simple geometric picture may be used for the classification of the waveforms measured for unknown gas species and gas mixtures.

FIG. 1
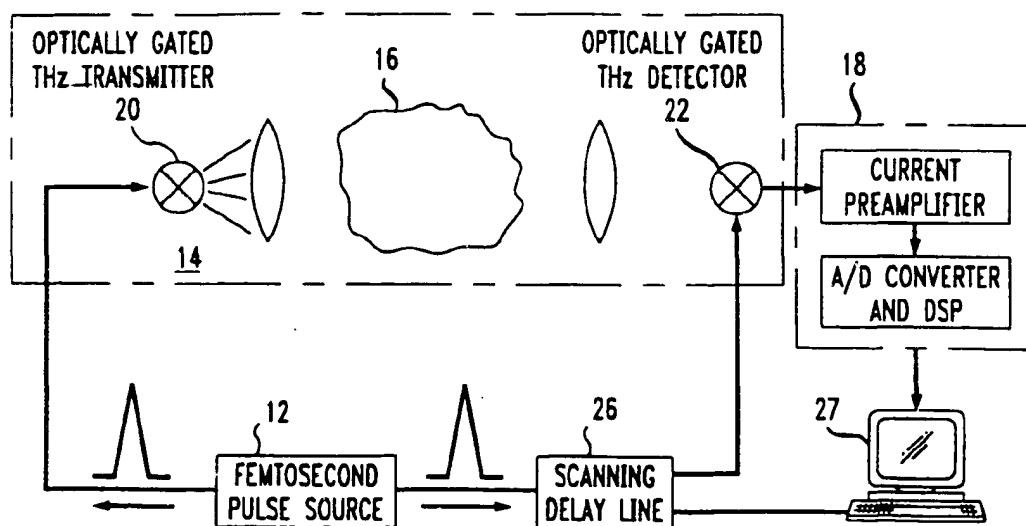
Description

Field of The Invention

The present invention relates generally to systems and methods of investigating various media or objects using reflected or transmitted radiation in the terahertz or far-infrared region of the spectrum and, more particularly, to the processing and analysis of terahertz waveforms.

Background of the Invention

The terahertz, or far-infrared region of the electromagnetic spectrum has some unique features. For example, THz waves easily penetrate most nonmetallic objects like paper, cardboard, plastics, and moderate thickness of many dielectrics, yet are absorbed by polar materials and liquids. Carriers in semiconductors show strong dielectric response in this region [M. van Exter and D. Grischkowsky, "carrier dynamics of electrons and holes in moderately doped silicon, " *Phys. Rev. B.,* vol. 41, pp. 840-842, (1992)], while metals are completely opaque to THz radiation. Polar gases such as water vapor, ammonia, HCl etc. have strong and very characteristic absorption lines in this region of the spectrum [H. Harde, S. Keiding, and D. Grischkowsky, *Phys. Rev. Lett.,* vol. 66, p. 1834 (1991)]. Polar liquids such as water and alcohols also absorb strongly in this frequency range [J. T. Kindt and C. A. Schmuttenmaer, "Far-infrared dielectric properties of polar liquids probed by femtosecond terahertz pulse spectroscopy," *J. Phys. Chem.,* vol. 100, 10373 (1996).] Consequently, the THz spectral range is becoming increasingly important for applications such as remote sensing of gases, quality control of plastic and composite materials, package inspection, and moisture analysis. These features can also be used for imaging in the THz frequency range [B. Hu and M. Nuss, "Imaging with terahertz waves, " *Opt. Lett.,* vol. 20, pp. 1716-18, (1995); M. Nuss, "Chemistry is right for T-ray imaging," *IEEE Circuits and Devices,* March 1996]. In addition, the terahertz frequency range has also been of considerable interest in spectroscopy. For example, the electronic properties of semiconductors and metals are greatly influenced by bound states (e.g., excitons and Cooper pairs) whose energies are resonant with THz photons [Nuss et al, "Terahertz time-domain measurement of the conductivity and superconducting bandgap in niobium," *J. Appl. Phys.,* vol. 70, pp. 2238-41, (1991)].

Despite its potential, the use of THz electromagnetic signals for spectroscopy and imaging has been hindered by a lack of suitable tools. For example, swept-frequency synthesizers for millimeter- and submillimeter-waves are limited to below roughly 100 GHz, with higher frequencies being heretofore available only through the use of discrete frequency sources. Fourier transform infrared spectroscopy (FTIR), on the other hand, remains hampered by the lack of brightness of incoherent sources. Additionally, FTIR methods are not useful since the real and imaginary part of response functions must typically be measured at each frequency. Finally, real-time imaging using the THz range of the electro-magnetic range has not been possible so far due to the poor sensitivity of detectors in this frequency range

In EP-A-0727671 a new spectroscopic imaging technique which overcomes the aforementioned deficiencies was disclosed.

This Terahertz ("T-ray") technique is based on electromagnetic transients generated opto-electronically with the help of ultrashort laser pulses (i.e., on the order of several femtoseconds (fs) or shorter). These THz transients are single-cycle bursts of electromagnetic radiation of typically less than 1 picosecond (ps) duration. Their spectral density typically spans the range from below 100 GHz to more than 5 THz. Optically gated detection allows direct measurement of the terahertz electric field with a time resolution of a fraction of a picosecond [Smith et al., *IEEE J. Quantum Electr.,* vol. 24, 255-260, 1988]. From this measurement, both the real and imaginary part of the dielectric function of a medium, which medium may be a solid, liquid, or gaseous composition, may be extracted in a rapid, straight-forward manner [M. Nuss and J. Orenstein, "Terahertz time-domain spectroscopy," in *Millimeter-wave spectroscopy of solids,* ed. George Gruener, Springer-Verlag, (Berlin 1997)]. The brightness of these THz transients exceeds that of conventional thermal sources, and the gated detection is several orders of magnitude more sensitive than bolometric detection.

There is a growing appreciation for the many potential commercial applications in which terahertz spectroscopy and imaging might be exploited [Nuss, "Chemistry is right for T-rays imaging," *IEEE Circuits and Devices,* March 1996, pp. 25-30]. Promising applications include industrial quality and process control, package inspection, moisture analysis, contamination measurements, chemical analysis, wafer characterization, remote sensing, and environmental sensing. A key ingredient to successful exploitation of any of the aforementioned applications, however, is a reliable, computationally practical technique for extracting the relevant spectroscopic information from the THz waveforms for a given application.

As will be readily appreciated by those skilled in the art, when a broad spectrum THz pulse is reflected by or transmitted through a medium under investigation, the acquired waveform contains a large number of data points. Previously, these waveforms have been analyzed by Fourier Analysis, which can be used to extract the frequency dependent absorption coefficient and refractive index of the material under investigation [M. Nuss et al, "Terahertz time-domain measurement of the conductivity and superconducting bandgap in niobium," *J. Appl. Phys.,* vol. 70, pp.

2238-41, (1991); and M. Nuss and J. Orenstein, "Terahertz time-domain spectroscopy," in *Millimeter-wave spectroscopy of solids,* ed. George Gruener, Springer-Verlag, (Berlin 1997)]. Although the frequency dependence of the absorption coefficient and refractive index over the large frequency range encompassed by the THz waveform (e.g., 100 GHz to a few THz) in many cases is unique from material to material, the amount of data which must be compared is too large to be practical for the various applications mentioned above. The inventors herein have thus identified the need for a method by which the relevant spectroscopic information contained in THz waveforms can be compressed to a much smaller set of data points without losing the relevant information contained in the original waveform. Such a compressed data set could be used, for example, to define a pixel in a T-ray image or in the recognition of materials or material compositions.

Summary of the Invention

For T-ray imaging, the wealth of information contained in the waveforms needs to be compressed so that the relevant information can be extracted, analyzed, and displayed in real time, for example with the help of a digital signal processor (DSP). In the present invention, a number of compression algorithms are devised that allow one to extract relevant information present in the THz waveform returned from a medium, and, optionally, to display the compressed data as pixels of a T-ray image of the medium. Furthermore, compression algorithms are devised that allow one to extract relevant information from the THz waveform returned form different media that are characteristic of the material. Illustratively, each such "classification" of the waveform may be associated with a material, and a code book of compressed waveforms may be compiled against which an unknown waveform can be compared. The objective of such a comparison step is to obtain compositional information about the medium under study.

Depending on the object under investigation, a number of compression techniques are presented in this invention that can be used to accomplish this task. Where only information about the transmissivity or reflectivity of the object is required (i.e., without spectral resolution) such as in package inspection, the compression technique may comprise an integration of the Fourier-transform of the waveform in a certain spectral range to obtain a measure of transmitted or reflected power, or a peak search to obtain the peak signal of the waveform. Because the spot-size for a diffraction-limited focal spot depends inversely on frequency, a partial integration over only the high-frequency portion of the THz frequency spectrum may be sufficient to increase the spatial resolution during T-ray imaging. Conversely, we have discovered that background absorption due to scattering can be largely avoided by integrating only over the low-frequency portion of the spectrum, because scattering increases with the forth power of frequency.

Since THz-TDS is a time-domain technique, another simple yet important compression step is timing extraction. Illustratively, timing extraction may be performed by finding the time-delay of the waveforms after passing through materials. This is useful, for example, in assessing thickness variations, or in determining the position of unknown objects in reflection geometry.

In addition to the simple peak-search method to find the temporal position of a THz waveform after transmission or reflection, the inventors have realized that typical THz waveforms have properties that are similar to mathematical wavefunctions called wavelets. Advantageously, a constant-scale wavelet analysis can be used to determine the temporal position of transmitted or reflected waveforms. Unlike the peak-position method, the wavelet analysis method of the present invention also works in the presence of multiple reflected waveforms, when the temporal position of many, isolated or overlapping THz waveforms has to be found.

In the case of sharp absorption features in the spectrum, such as after passage of THz waves through polar gases, the inventors herein have observed that the action of a resonantly excited gas or gas mixture may be modeled in the time domain by a linear digital filter that reshapes an incident THz waveform to produce a corresponding free-induction decay (FID). According to this aspect of the invention, the free induction decay exhibited by gases excited by far-infrared (terahertz) pulses is analyzed using digital signal processing techniques. These circumstances strongly resemble the basic problems of spectral estimation for phonetic recognition, such that theory and algorithms from speech recognition systems are applicable. Illustratively, a correlation type of analysis known as *linear predictive coding* (LPC) may be employed to extract and parameterize the spectral features of a waveform. A simple geometric picture enables the use of these parameters for classification of individual gas species as well as quantitative mixture analysis where more than a single gas species is present.

In more general situations, the THz waveforms are modified in both amplitude and phase by transmission through or reflection by the medium under investigation. In this case, more complex signal processing has to be performed to compress the data and extract the relevant information. As noted above, the inventors have observed that the action of a medium can be modeled in the time domain by a linear digital filter that reshapes an incident THz waveform to produce a corresponding output waveform. A very general way of compressing such a linear filter is by use of the digital filter bank method. The inventors have also realized that, in many cases, the mathematical algorithms used for speech processing and recognition can be modified to perform THz waveform processing and recognition.

Brief Description of the Drawings

The features and benefits of the inventions will be better understood from a consideration of the detailed description which follows taken in conjunction with the accompanying drawings, in which:

FIG. 1 depicts a THz imaging system employing THz radiation to achieve chemical recognition in accordance with the present invention;

FIG. 2A shows a typical time-domain THz waveform obtained in a system of FIG. 1 and FIG. 2B depicts its power spectrum obtained by a Fourier transform (FFT) of the time-domain waveform;

FIGS. 3A-3E schematically show how THz waveforms may be altered by transmission or reflection from different materials or objects. From top to bottom: 3A) uniform attenuation, 3B) high-frequency absorption, 3C) frequency-dependent index, 3D) thickness variations, and 3E) sharp absorption lines;

FIGS. 4A-4C show T-ray images of a resolution target. In each case, the digital signal processor performs a FFT of the waveform and integrates over portions of the magnitude of the FFT. In FIG. 4A, integration over the frequency portion 0.1-0.3 THz of the spectrum is performed; in FIG. 4B, integration over the frequency portion 0.3-0.5 THz of the spectrum is performed; and FIG. 4C depicts integration over the 0.5-0.7 THz frequency portion of the spectrum;

FIGS. 5A and 5B depict T-ray images of strands of human hair, without (right hand side) and with (left hand side) moisturizer applied. The images in FIG. 5A are obtained by integrating the power spectrum over a 0.25-0.5 THz range, while FIG. 5B depicts that use of integration over the 1.25-1.5 THz frequency range;

FIG. 6 shows a T-ray image of a flame, obtained by measuring the difference in arrival time of the focused THz pulses through the flame;

FIG. 7 is a THz waveform for HCl vapor at 13 kPa and a reference waveform measured for an evacuated gas cell, the HCl waveform being shifted by 1.5 nA for clarity; FIG. 8 is a power spectrum estimate, using H_2O at a pressure of 2.8 kPa, from an LPC analysis of order M = 50 (thick solid line) and from a Fourier analysis (dashed line) using zero padding techniques to construct an input array of 8192 data points, the thin solid and dotted lines representing the integrated power spectral density for the LPC and Fourier analysis, respectively;

FIG. 9 is a two dimensional LPC vector representation of a binary mixture evaluated in accordance with the technique of the present invention , in which **a** is the vector sum of the coded vectors for the individual species (a_i) weighted by their mole fractions x_i, in which $\{b_j\}$ is an orthogonal basis used to calculate the orthogonal projections y_i, and in which a linear transformation establishes the relation between the two bases $\{a_i\}$ and $\{b_j\}$; and

FIG. 10 is a graphical representation of chemical recognition, in accordance with the present invention, of mixtures of NH_3 and H_2O, in which each mixture is represented by a single point representing the estimated partial pressure of the two gas species.

Detailed Description of The Invention

In accordance with the present invention, chemical recognition utilizing THz waveforms is performed using an imaging apparatus such as the one disclosed in EP-A-0727671.

Such an apparatus, indicated generally at 10, is shown in FIG. 1. As can be seen in FIG. 1, THz imaging apparatus 10 includes a source 12 of repetitive optical pulses of femtosecond duration, imaging arrangement 14 by which THz radiation is generated, directed at a medium under investigation 16, and detected upon transmission through or reflection by the medium, and analysis circuitry indicated generally at 18. Source 12 may be configured, for example, as a solid state laser like the Ti:Sapphire laser, which has a wavelength near 800 nm and a typical repetition rate of about 100 MHz. Alternatively, source 12 may be configured as a femtosecond Erbium-Doped Fiber Laser operating at a wavelength near 1.5 μm. In the illustrative embodiment depicted in FIG. 1, imaging arrangement 14 includes an optically gated THz transmitter 20 and an optically gated THz detector 22. A beam splitter (not shown) divides the output of source 12 into two beams, the pulses of which are used to optically gate transmitter 20 and detector 22. A variable delay line 26, varies the optical delay between the respective gating pulses to acquire a sampled replica of the THz waveform.

FIG. 2A shows an exemplary THz waveform obtained by the setup depicted in FIG. 1, and FIG 2B shows the power spectrum of the THz waveform obtained by computing the magnification of the Fourier transform of the waveform in FIG. 2A.

As discussed in the aforementioned co-pending application, THz imaging apparatus 10 may be employed to analyze temporal distortions of the THz waveforms returned from an object for each pixel in real time to obtain information about the composition or properties of the object at the area illuminated by the THz beam. FIGS. 3A-3E show schematically how THz waveforms may be altered by transmission through an object with different properties. Although this is illustrated by using the example of a transmission geometry, it can easily be appreciated by someone skilled in the art that similar distortions of THz waveforms may occur in reflection. In FIG. 3A, the THz waveform is going through a material with frequency-independent absorption and refractive index, resulting in a returned waveform that is an attenuated replica of the input waveform. As shown in FIG. 3B, high-frequency absorption of the traversed medium leads to selective attenuation of the high-frequency portion of the waveform and consequent broadening of the waveform in the time-domain. In FIG. 3C, the material with frequency-dependent refractive index leads to a "chirped" waveform, in which different frequency components of the waveform travel at different phase velocity, thus leading to a time-dependent frequency variation within the pulse. As shown in FIG. 3D, thickness or index variations of the sample lead to differences in arrival time of the waveform, depending on which portion of the object was traversed by the beam. As shown in FIG. 3E, sharp absorption lines within the spectrum of the THz pulse, such as for polar gases such as water vapor, CO_2, HCl vapor, etc., lead to oscillations at the frequencies of the sharp lines in the trailing part of the pulse.

It is known in the field of THz spectroscopy that information about the complex dielectric constant of the material under investigation can be obtained by comparing the frequency-dependent amplitude and phase of the input and returned waveforms. Heretofore, this has been accomplished by computing the complex Fourier transform of both input and returned waveforms, and dividing the complex Fourier spectrum of the returned waveform by the complex Fourier spectrum of the input spectrum. The complex dielectric constant (absorption coefficient and refractive index) can then be obtained by taking a log of the waveform and dividing the result by the thickness of the medium [Nuss, Orenstein, 1997]. Since many materials have characteristic variations in the frequency dependence of their absorption constant and refractive index over the 100 Hz to few THz frequency range covered by the THz waveform, a comparison of the Fourier spectra of the returned THz waveforms can give an indication of and in some cases uniquely identifies the composition of the object.

For T-ray imaging, the wealth of information contained in the waveforms needs to be compressed so that the relevant information can be extracted, analyzed, and displayed in real time, for example with the help of a digital signal processor (DSP). Depending on the object under investigation, a number of compression techniques are presented in this invention that can be used to accomplish this task.

Where only information about the transmissivity or reflectivity of the object is required without spectral resolution such as in package inspection, compression can comprise an integration of the Fourier-transform of the waveform in a certain spectral range to obtain a measure of transmitted or reflected power, or a peak search to obtain the peak signal of the waveform. Where the optical system can focus to a diffraction-limited focal spot, as described in European Patent Application No. 97306745.7 the spot-size will depend inversely on frequency. Advantageously, only a partial integration - over only the high-frequency portion of the THz frequency spectrum - is needed to increase the spatial resolution during T-ray imaging. As an illustration, illustrative FIGS. 4A and 4C show T-ray images of a resolution target over three frequency ranges wherein the T-ray image has been obtained by integrating the THz spectrum over three different frequency ranges (0.1-0.3, 0.3-0.5, and 0.5-0.7, respectively). In each case, the digital signal processor performs a FFT of the waveform and integrates over portions of the magnitude of the FFT. Because the of the capability of the system to focus to a diffraction limited spot, i.e., with a spot size that varies inversely proportional to the frequency, integration over the high-frequency portion of the power spectrum leads to images with better spatial resolution.

We have further discovered that background absorption due to scattering can be largely avoided by integrating only over the low-frequency portion of the spectrum, because scattering increases with the forth power of frequency. Scattering can be particularly severe when the size of the scattering centers approaches the wavelength of the THz radiation, which is about 300 mm for a 1 THz center frequency. FIGS. 5A and 5B depict T-ray images of strands of human hair, untreated, and with treated moisturizer applied. The images in FIG. 5A are obtained by integrating the power spectrum over a 0.25-0.5 THz range, while those in FIG. 5B depict results obtained by integrating over the 1.25-1.5 THz frequency range. In the T-ray images obtained using high-frequency components of the spectrum, it is difficult to see the added absorption due to the additional moisture on the left hand side because of the strong background absorption due to scattering. On the other hand, a T-ray image obtained by integrating over only the low-frequency components of the power spectrum clearly shows the difference due to the presence of moisture in the hair.

Because THz-TDS is a time-domain technique, another simple yet important compression step is timing extraction, for example by finding the time-delay of the waveforms after passing through materials. This is useful for example in assessing thickness variations, or in determining the position of unknown objects in reflection geometry. As an example, FIG. 6 shows a gray scale T-ray image of a flame, obtained by measuring the difference in arrival time of the focused

THz beams through the flame. The DSP was instructed in this case to determine the peak position of the waveform in time and apply a gray scale to the temporal shift of the waveform. Each contour corresponds to a 5 femtosecond incremental shift in time delay. It will be appreciated from FIG. 6 that, although the duration of the THz pulse is on the order of 1 ps, the position of the peak of the waveform can be determined with an accuracy approaching a few femtoseconds.

In addition to the simple peak-search method to find the temporal position of a THz waveform after transmission or reflection, the inventors have observed that typical THz waveforms have properties that are similar to mathematical wavefunctions called wavelets [M. Vetterli and J. Kovacevic, "Wavelets and Subband Coding," Prentice-Hall, (1995) reference]. Therefore, the inventors have recognized that a constant-scale wavelet analysis can conveniently be used to determine the temporal position of transmitted or reflected waveforms. Unlike the peak-position method, the wavelet analysis method also works in the presence of multiple reflected waveforms, when the temporal position of many, isolated or overlapping THz waveforms has to be found.

In the case of sharp absorption features in the spectrum, such as after passage of THz waves through polar gases, the inventors have identified that the action of a resonantly excited gas or gas mixture can be modeled in the time domain by a linear digital filter that reshapes an incident THz waveform to produce a corresponding free-induction decay (FID). Extraction and analysis of the spectral features of the modeled waveform may be performed using techniques similar to those employed for phonetic recognition. By way of illustrative example, a correlation type of analysis known *as linear predictive coding* (LPC) may be used to extract and parameterize the spectral features of a waveform. Linear prediction is an efficient model for analyzing sharp spectral features. The waveform is treated as a result of an all-pole filtering of a noise input and the filter coefficients are the parameters of the LPC. Ideally, all the spectral information is transferred to the filter coefficients and the input becomes white noise.

In accordance with the present invention, the measured waveform is expressed as a linear combination (i.e. a weighted sum) of its past M values. The waveform is denoted $s(n)$, and is measured at points equally spaced in time. By including an error term, $e(n)$, the sampled waveform can be expressed:

$$s(n) = \sum_{k=1}^{M} a_k s(n-k) + e(n), \qquad \qquad 1$$

where e(n) is the residual that accounts for the equality to hold, and a_k are the parameters of the linear prediction. As will be readily appreciated by those skilled in the art, an LPC model can be described in both the time, t, and the frequency domain f. In the interest of clarity and simplicity of illustration, only a frequency formulation will be described in detail herein. It should, however, be emphasized that the experimental implementation presented below is based directly on a time-domain analysis and does not involve Fourier-transformations .

To illustrate the basic physical concepts involved, the z-transform of Eq. (1) is presented:

$$S(z) = \sum_{k=1}^{M} a_k z^{-k} S(z) + E(z), \qquad \qquad 2$$

where z = exp(i2πfΔ) and Δ-1 is the sampling frequency. One finds that the signal, $S(z)$, is constructed from the residual, $E(z)$, by an *all-pole* filtering:

$$S(z) = H(z)E(z), \qquad \qquad 3$$

where

$$H(z) = \frac{1}{1 - \sum_{k=1}^{M} a_k z^{-k}}. \qquad \qquad 4$$

The LPC coefficients, ak, contain the "information" of the spectral content of the waveform. The advantages of using LPC for chemical recognition of gases are threefold. First, poles provide an accurate representation for an un-

derlying power spectrum that has sharp spectral lines. This is in contrast to Fourier analysis where a signal waveform is expanded into a Fourier series. Such series can have only zeros, not poles, and must attempt to fit sharp spectral features with a polynomial which requires a large number of coefficients. Second, by using LPC one can always rely on the same number of coefficients, M, regardless of the number of resonances excited. Accordingly, simple geometric pictures can be used instead of complicated template matching techniques for the classification of waveforms. Finally, the algorithms for LPC analysis may be executed quickly and the implementation is suitable for parallel processing in contrast to Fourier analysis where the entire waveform has to be sampled before further processing.

The optimum values for the LPC coefficients are calculated from least square principles, that is, it is desired to minimize the total squared residual: $\varepsilon = \Sigma e(n)^2$. The analysis yields a system of normal equations that relates the LPC coefficients to values of the autocorrelation function, $\Sigma k\, s(k)s(k-n)$, of the time series representing the waveform. The problem can be solved in terms of numerical matrix inversion schemes such as the Levinson-Durbin algorithm or by using Burg's algorithm for maximum entropy spectral analysis. For a detailed discussion of these algorithms and of their application, reference may be had to a book entitled *Digital Signal Processing, Principles, Algorithms, and Applications*, by J. G. Proakis and D. G. Manolakis, Prentice Hall, New Jersey (1993).

In the previous examples, compression algorithms for THz waveforms have been presented that extract certain relevant information from the waveform. The compressed data set can now either be displayed to produce a T-ray image, or can be further processed for the purpose of extracting compositional information about the medium under investigation. In the context of the present invention, the compressed data set is treated as a vector in a high dimensional vector space. If the vector represents a known species, it can be stored in a code book for later use in a recognition procedure -- with each material of interest corresponding to one vector in the code book. Accordingly, when it is desired to recognize an unknown species, one may measure the THz waveform, extract the vector parameters, and perform a comparison with the known species from the code book. In the following, this procedure is illustrated using the example of gas recognition and Linear Predictive Coding (LPC) as compression and classification procedure:

In accordance with an illustrative embodiment of the present invention, the parameters from an LPC analysis are treated as a vector, $a = (a_1, a_2, \ldots a_M)$, in an M-dimensional vector space. Vectors representing known gas species are stored in a codebook, and by successive measurements and subsequent LPC analysis for a variety of different gases, the codebook is built. The final codebook contains a set of linearly independent vectors $\{a_i\}, i = 1,2,\ldots,p$ and $p < M$, that span a subspace of the M-dimensional vector space. A single-species-recognition procedure consists in a full search through the codebook to find the "best" match between the trial vector, i.e., the LPC coefficients of an unknown species, and the vectors in the codebook. This is accomplished, for example, by calculating the Euclidean distances between the trial vector and the coded vectors and use the minimum distance as classification criteria.

Gas mixture analysis according to the present invention is made possible by the fact that the LPC vectors from different gases add in the M-dimensional vector space. Hence, a coordinate representation, x_i, of the trial vector in the $\{a_i\}$-basis becomes a measure of the mole fractions of the species present in the mixture. These principles are illustrated in FIG. 9 using a binary mixture as an example. The LPC vector representing the mixture, a, is the vector sum of the coded vectors for the individual species (a_1 and a_2) weighted by their mole fractions. In practice, an orthogonal basis, $\{b_i\}$, is constructed and is used to calculate the orthogonal projections y_i. A linear transformation establishes the relation between the two bases $\{a_i\}$ and $\{b_i\}$

Chemical recognition in accordance with the present invention was evaluated experimentally by using the arrangement 10 of FIG. 1 to obtain sampled replicas of THz waveforms. Four gases, HCl, NH_3, H_2O and CH_3CN, in a pressure range of from 0.3 to 13 kPa, were evaluated by placing them into a gas cell 30.5 cm in length. In each experiment, a collimated THz beam was transmitted through the gas cell.

With particular reference to FIG. 7, there is shown an example of a THz waveform resulting from propagation in an HCl atmosphere at a pressure of 13 kPa. The measured FID shows fast oscillations due to the frequency beating between resonances at 0.626, 1.251 and 1.876 THz. [H. M. Pickett, R. L. Poynter, and E. A. Cohen, *Submillimeter, Millimeter, and Microwave Spectral Line Catalog*, accessed via World Wide Web (http://spec.jpl.nasa.gov) from the Jet Propulsion Laboratory, Pasadena CA]. Also shown in FIG. 7 is a reference waveform measured using an evacuated gas cell. The waveforms were obtained by sampling N = 1024 points with a temporal resolution of $\Delta = 93$ fs.

Prior to the LPC analysis, the measured waveforms were preprocessed using finite impulse response filtering and time-domain windowing. That is, the sampled waveform was convoluted with the impulse response of a bandpass filter and is multiplied by window functions. This is performed to stabilize the LPC by reducing the risk of predicting spurious frequency peaks and eliminating frequencies beyond the bandwidth of the THz pulse. Because only the coherent irradiance from the gas is of interest (not the spectrum of the source), LPC analysis commences with values appearing after the impulse excitation.

The LPC coefficients obtained in the manner described above can be used for a power spectrum estimation of the input waveform. This technique is known as the maximum entropy method (MEM) and is described in the text by Rabiner and Juang referenced above. FIG. 8 shows a comparison between the maximum-entropy-method (MEM) power spectrum, calculated from the LPC coefficients (solid line) and the Fourier-transform power spectrum (dashed

line) for a waveform that has been propagated through H_2O vapor at 2.8 kPa. The MEM spectrum is estimated from the power spectrum of the all-pole filter, $|S(z)|^2 = a_o|H(z)|^2$, where $a_o = min(\varepsilon)$ is the minimum squared residual. The comparison between the two spectra, while not used for recognition purposes, is useful as a confidence measure of the LPC. The frequency resolution obtained from the Fourier analysis is better that the resolution obtained from the
5 LPC. However, the LPC analysis correctly predicts the power, i.e., the integral of the power spectral density, in an absorption line. For comparison, the inset on the figure shows the magnitude of the Fourier-transform of our source.

A gas mixture determination in accordance with the present invention relies on the fact that the LPC vectors corresponding to different gases add. Hence, a coordinate representation, x_i, of the trial vector in the $\{a_i\}$-basis becomes a measure of the mole fractions of the species present in the mixture. j These principles are illustrated in FIG. 9 using
10 a binary mixture as an example. The LPC vector representing the mixture, a, is a vector sum of the coded vectors for the individual specias (a_1 and a_2) weighted by their mole fractions. In practice, an othrogonal basis, $\{b_i\}$, is constructed and is used to calculate the orthogonal projections y_i. A linear transformation establishes the relation between the two bases $\{a_i\}$ and $\{b_i\}$. From the projections of the resulting vector onto the vectors in the codebook, one can determine the mole fractions of the various species. To demonstrate the use of the chemical recognition system for mixture analysis
15 a binary mixture of NH_3 and H_2O was evaluated. The result is shown in FIG. 10.

The mixtures were prepared starting with NH_3 vapor at 12.9 kPa and successively exchanging with H_2O vapor in increments of a few percent per volume. For each mixture, a THz waveform was recorded. From the LPC analysis and the geometric interpretation of the LPC vector, mole fractions of the two species were calculated. The mole fractions are converted into partial pressures by scaling to the pressure of the corresponding coded vectors. Nonvanishing
20 projections onto species in the codebook, which were not present in the mixture (HCl and CH_3CN), are attributable to experimental uncertainties and are used to estimate error bars. Error bars are only displayed when they become larger than the actual data point.

From FIG. 10, we find that the chemical recognition system is capable to trace the transition from NH_3-dominated mixtures (lower right) to the H_2O-dominated mixtures (upper left). When a mixture is strongly dominated by NH_3 the
25 prediction yields a slightly negative pressure for H_2O and a pressure for NH_3 that is larger than the starting pressure. Obviously, this situation does not have any physical meaning, and results from insufficient statistics for the coded vectors. By using clustering techniques of vectors obtained from repeated measurements on a particular gas species we should be able to account for this inaccuracy. One clustering techniques which is suitable for this purpose is disclosed in a text by L. Rabiner and B. -H. Juang entitled *Fundamentals of Speech Recognition* (Prentice Hall, 1993).
30 The algorithms used in accordance with the illustrative embodiment of the present invention discussed in detail above have been chosen to be suitable for real time application. The recent progress in digital signal processing (DSP) technologies allows such an approach to be implemented in an easy and inexpensive manner. It should, however, be readily appreciated by those skilled in the art that although a DSP implementation is preferred, discrete filters may also be employed to obtain the filtering coefficients used in the illustrative embodiment described above using LPC and, in
35 fact, that alternate classification algorithms may be utilized instead of LPC. It will therefore be recognized that the embodiments shown and described in detail herein are intended to be merely illustrative of the inventive concepts involved.

40 **Claims**

1. A method of processing and analyzing Terahertz waveforms returned by a medium having at least one constituent, comprising the steps of:

45   receiving terahertz waveforms that are one of transmitted through or reflected by the medium;

      obtaining a digitally sampled replica of the received terahertz waveforms; compressing the digitally sampled replicas into respective data sets of reduced size;

50   displaying said reduced size data sets as pixels of an image of the medium, each pixel being representative of the terahertz waveform returned from a distinct portion of the medium illuminated by an incident terahertz pulse.

2. The method of claim 1, wherein said compressing step includes:
55   integrating portions of a power spectrum obtained from a FFT of the returned waveform to thereby obtain a measurement of received power proportional to the power within said portions of the spectrum.

3. The method of claim 2, wherein said step of integrating comprises integrating higher frequency portions of the

obtained power spectrum to thereby obtain enhanced spatial resolution, and

 integrating lower frequency portions of the obtained power spectrum to thereby remove scattering background.

4. The method of claim 1, wherein said step of compressing includes at least one of: (a) determining a peak amplitude of the returned waveform to thereby obtain an approximation of the received power, and (b) determining whether or not the amplitude of the returned waveform exceeds a particular threshold to thereby determine the presence or absence of an object.

5. The method of claim 1, further including a step of measuring the arrival time of the returned waveform to thereby ascertain a temporal delay of the terahertz waveform following transmission through or reflection by the medium, and, when the refractive index of the medium is substantially constant, a step of using temporal delay information to identify variations in the thickness of an object at multiple, spatially distinct, locations traversed by terahertz waveforms, and, when the thickness of the medium is substantially constant, a step of using temporal delay information to identify variations in the refractive index of an object at multiple, spatially distinct locations traversed by terahertz waveforms.

6. The method of claim 2, further including, when the thickness of the medium is substantially constant, a step of using the obtained power measurement to identify variations in the absorption coefficient of the medium at multiple, spatially distinct locations traversed by terahertz waveforms, and using the obtained power measurement to identify variations in the reflection coefficient of the medium at multiple, spatially distinct locations.

7. The method of claim 5, wherein the arrival time is measured by at lease one of: (a) locating a peak value of the waveform, and (b) by performing a wavelet analysis on the returned waveform.

8. A method of processing and analyzing Terahertz waveforms returned by a medium having at least one constituent, comprising the steps of

 receiving a terahertz waveform that is one of transmitted through or reflected by the medium;

 obtaining a digitally sampled replica of the received terahertz waveform;

 compressing the digitally sampled replica into a data set of reduced size representing a classification of the waveform; and

 associating the data set with the at least one constituent to thereby identify said at least one constituent.

9. The method according to claim 8, wherein the classification uses linear predictive coding of the returned waveform to obtain a set of coefficients reproducing said data set of reduced size.

10. The method of claim 8, wherein the classification utilizes at least one of (a) a filter bank parameterization of the returned waveform, and (b) a wavelet analysis of the returned waveform.

11. The method according to claim 8, wherein the associating step comprises comparing the waveform classification obtained during the compressing step with waveform classifications obtained from media of known composition.

12. The method of claim 8, wherein the compressing step includes at least one of : (a) forming a vector comprising coefficients of the reduced size data set, and (b) forming, for media of known composition, ortho-normal vectors comprising coefficients representative of at least some of the known media; and storing the ortho-normal vectors in memory for subsequent comparison with test vectors returned from media under investigation.

13. The method of claim 12, wherein said associating step includes performing a measurement of the Euclidean distance between the vector associated with the medium under investigation and at least one of the ortho-normal vectors stored in memory; and

 selecting, on the basis of the distance measurement step, an ortho-normal vector closest to the vector formed for the medium under investigation; and

identifying, as a constituent of the medium under investigation, the known composition corresponding to the selected ortho-normal vector.

14. The method of claim 13, wherein the test vector is within an error distribution identical to one associated with the selected ortho-normal vector, wherein the test vector is constructed as a linear combination of at least two stored ortho-normal vectors, and wherein the coefficients of the linear combination indicate the relative concentration of constituents within the medium under investigation.

15. A method of processing and analyzing temporal amplitude and phase differences between terahertz waveforms incident on and terahertz waveforms returned by a medium having at least one constituent, comprising the steps of:

obtaining digitally sampled replicas of the incident and received terahertz waveforms;

expressing the returned waveform as a linear filtering operation of the incident waveform, the filter coefficients employed in the linear filtering operation representing a classification of the at least one constituent;

associating the filter coefficients with the at least one constituent to thereby identify said at least one constituent.

16. The method of claim 15, wherein the linear filtering operation is performed using at least one of : (a) a digital signal processor, and (b) a bank of discrete bandpass filters.

17. The method of claim 15, wherein the linear filtering coefficients are found by calculating the ratio of Fourier transforms of the returned and incident waveforms and Fourier transforming said ratio into the time domain.

18. The method of claim 15, wherein the linear filtering coefficients are found by performing a least squares error analysis of the incident and returned waveforms.

19. The method of claim 15, further including storing, for media of known composition, linear filtering coefficients representative of at least some of the known media.
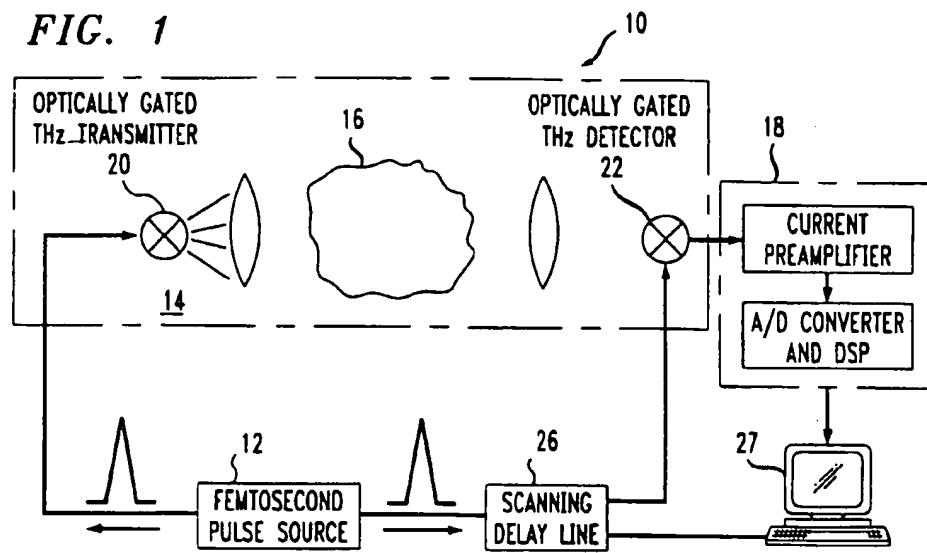
FIG. 1

FIG. 2A
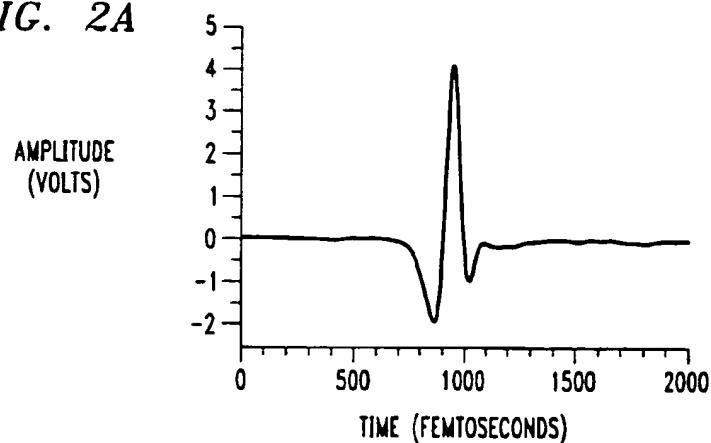
FIG. 2B
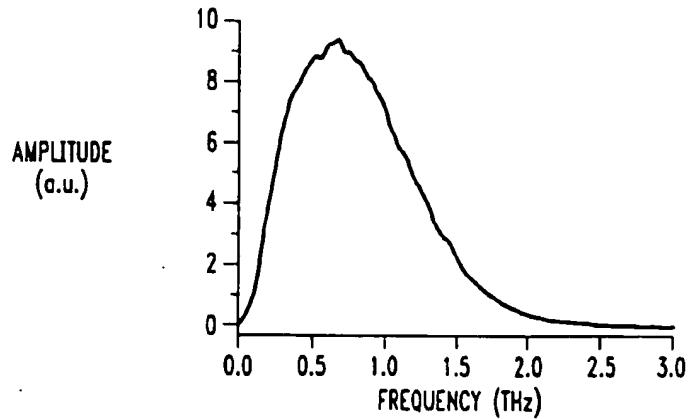
INPUT THz
WAVEFORM

OUTPUT THz
WAVEFORM

FIG. 3A

TIME

MATERIAL

FIG. 3B

FIG. 3C

FIG. 3D

FIG. 3E

NEWPORT RESOLUTION TARGET: METAL ON 1.6mm BK7

FIG. 4A

Y-POSITION
(mm)

0.1-0.3 THz

X-POSITION (mm)

FIG. 4B

Y-POSITION
(mm)

0.3-0.5 THz

X-POSITION (mm)

FIG. 4C

Y-POSITION
(mm)

0.5-0.7 THz

X-POSITION (mm)

TERAHERTZ IMAGES OF STRANDS OF HAIR

FIG. 5A

LOW FREQUENCY FILTER
(0.25 - 0.5 THz)

POSITION
(mm)

80-85% TRANSMISSION
75-80
70-75
65-70
60-65
55-60
50-55

FIG. 5B

HIGH FREQUENCY FILTER
(1.25 - 1.5 THz)

POSITION
(mm)

80-85% TRANSMISSION
75-80
70-75
65-70
60-65
55-60
50-55

FIG. 6

POSITION
(mm)

POSITION
(mm)
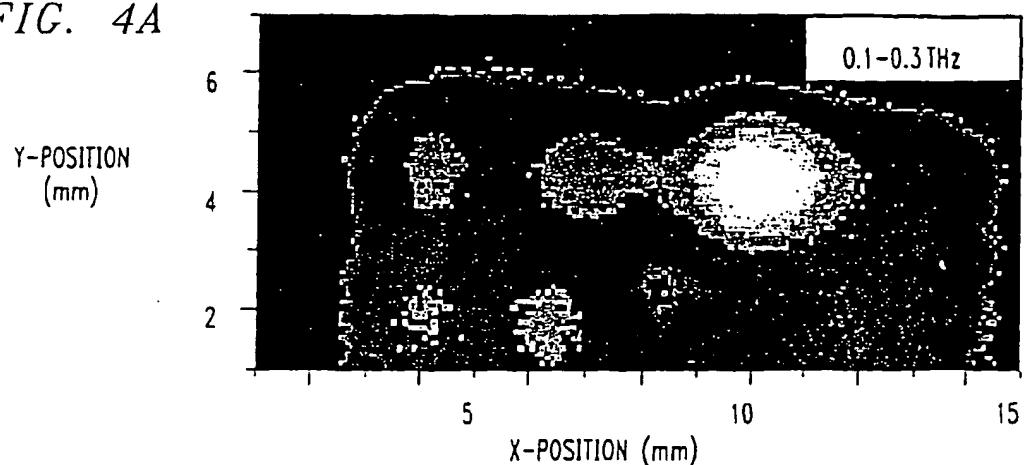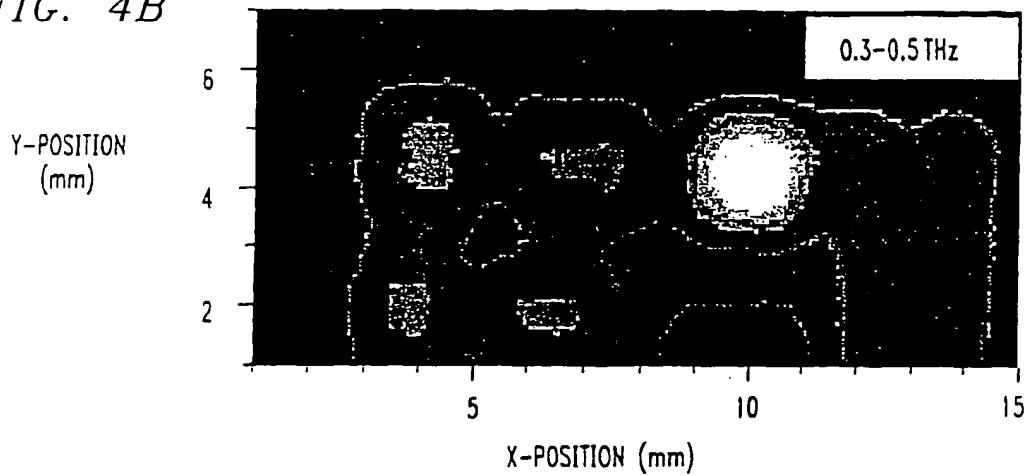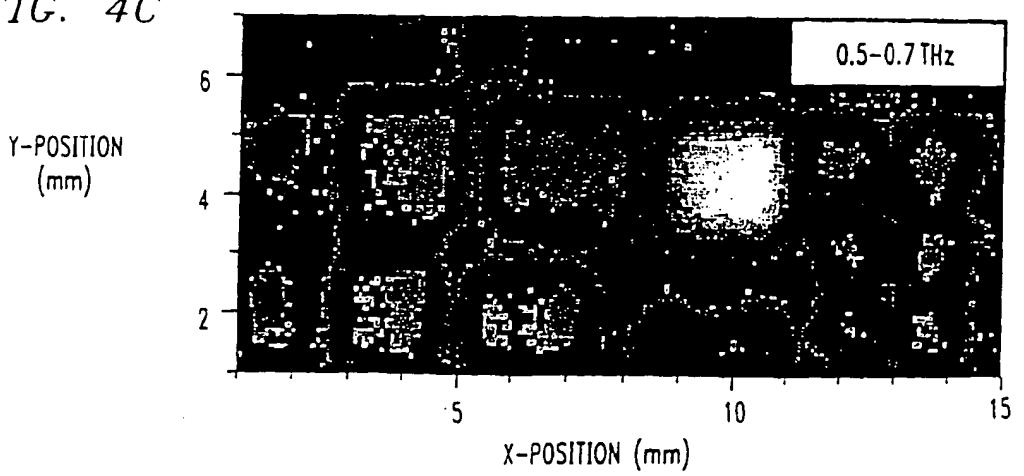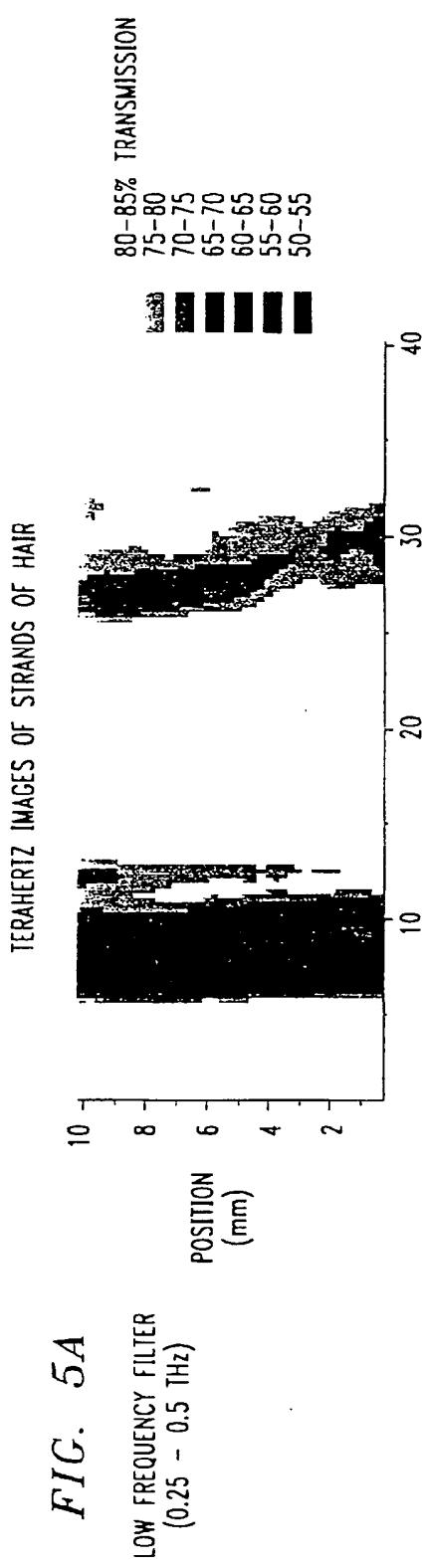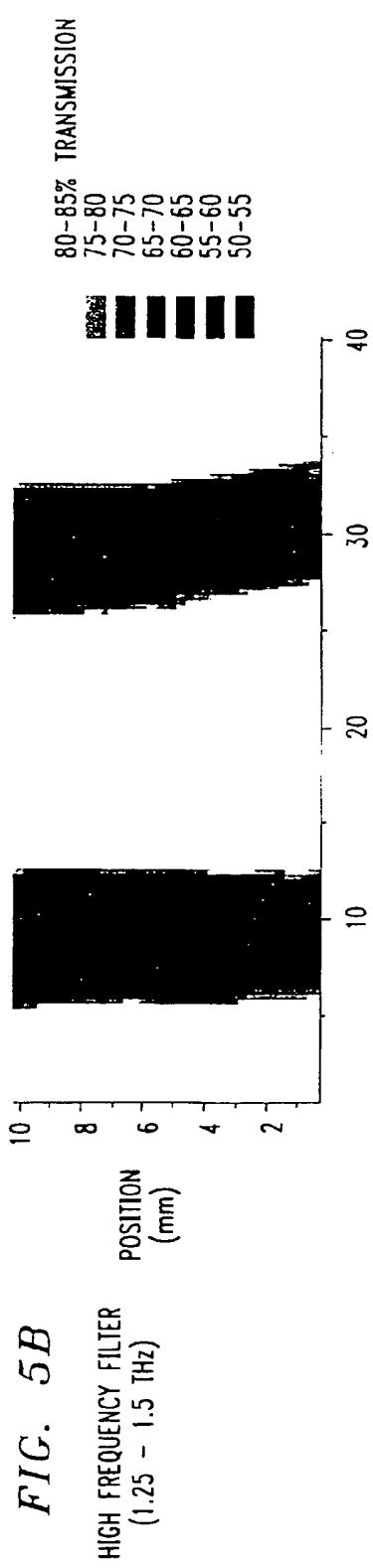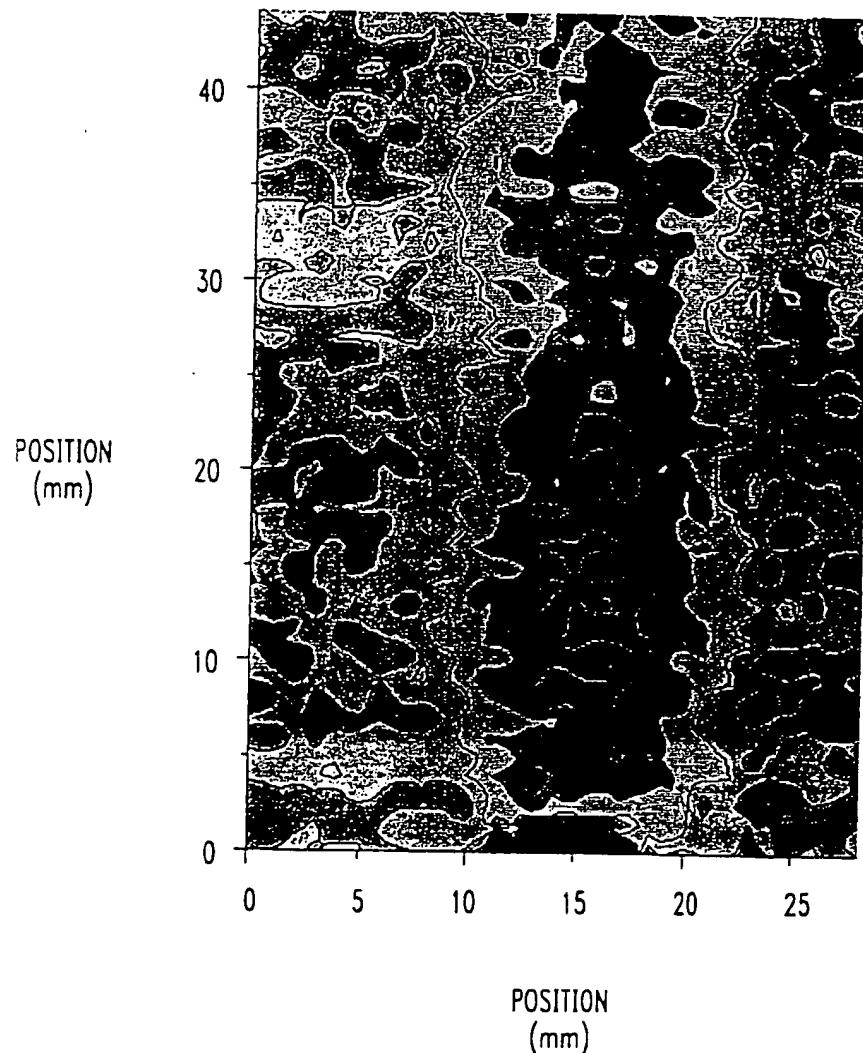
FIG. 7
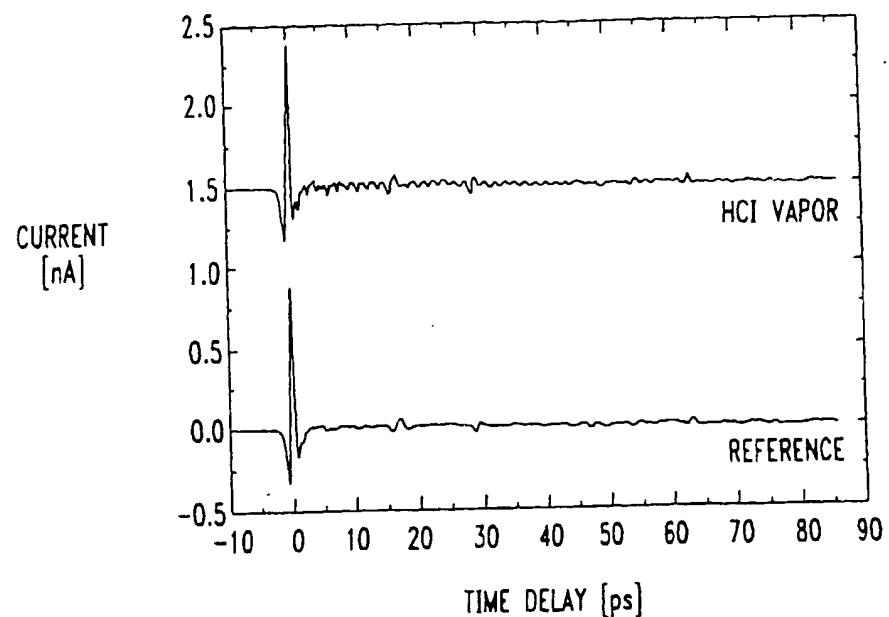
FIG. 8
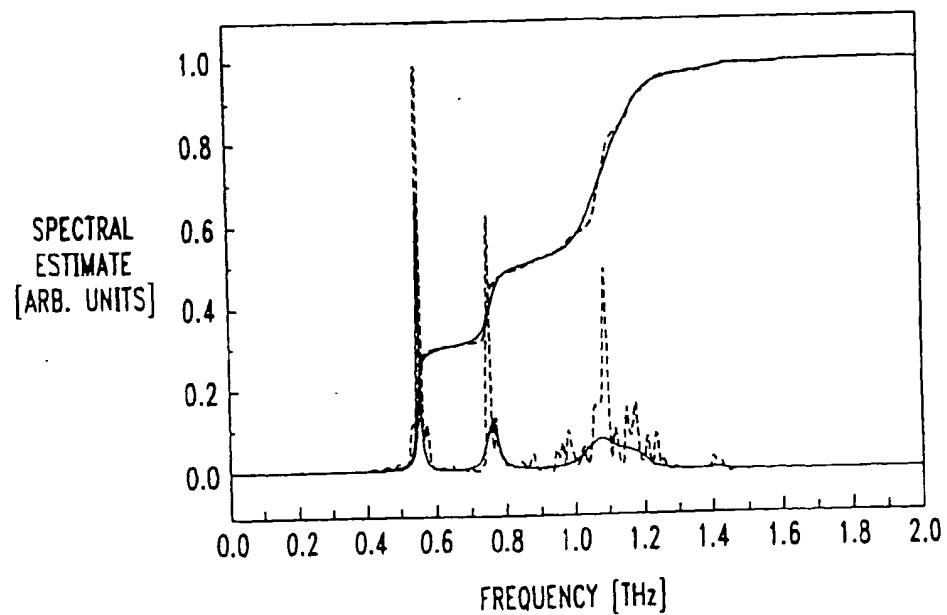
FIG. 9
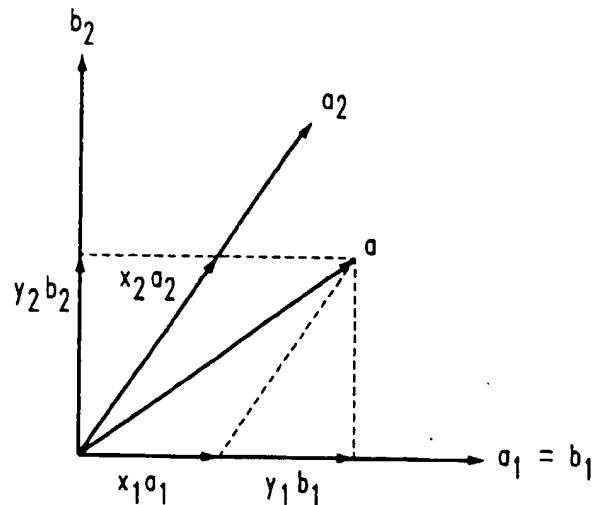
FIG. 10

THIS PAGE BLANK (USPTO)